Feature extraction/data compression

C. Bohm and K Hultkvist

Stockholm University

Overview
Simple S-18 algorithms
Performance evaluation
Noise modelling for simulations
Conclusions

The task of the Feature Extractor (FEX) is:
to reduce the bandwidth by recognizing single photons (SPE),
transmitting only their parameters –
time and amplitude

-**Efficient methods exists but**
-**Limited computational resources (FPGA) require**
-**Resonable compromizes**
-**Especially for the string 18 DOMs**

SPE detction errors:
False positive
Will pass a MPE as a SPE
Serious error – information loss
Must be minimized – 0%

False negative:
Transmission of SPE as waveform
Wastes bandwidth – not fatal – minimize

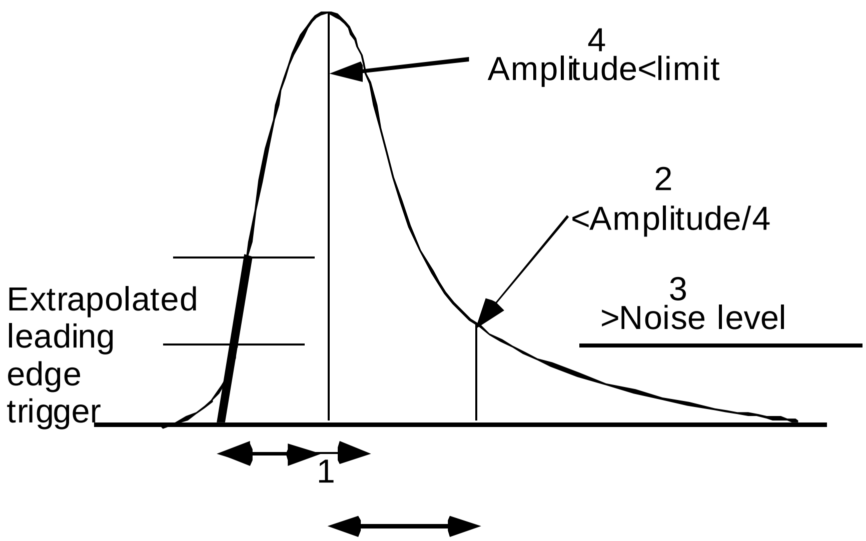
Reduce false positive at expense of accepting more false negative

SPE measurement errors:
Timing precision about 1 ns
Amplitude precision – 5%?:

Fraction of events containing one SPE – ~10-15%

String 18 algorithms

Reduced FPGA resources
(Altera 10k50 - 85 % used for other purposes
In icecube Alterra EPXA1 – twice as large)
Simplified approach



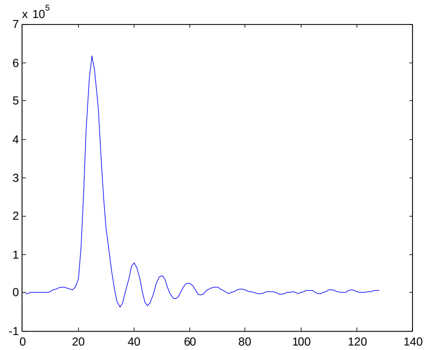
We will have full read-outs of prescaled SPEs

Base line tracking (suggested by D. N.)
Histogram samples
If base line below k% level increment base line one unit
Else decrement

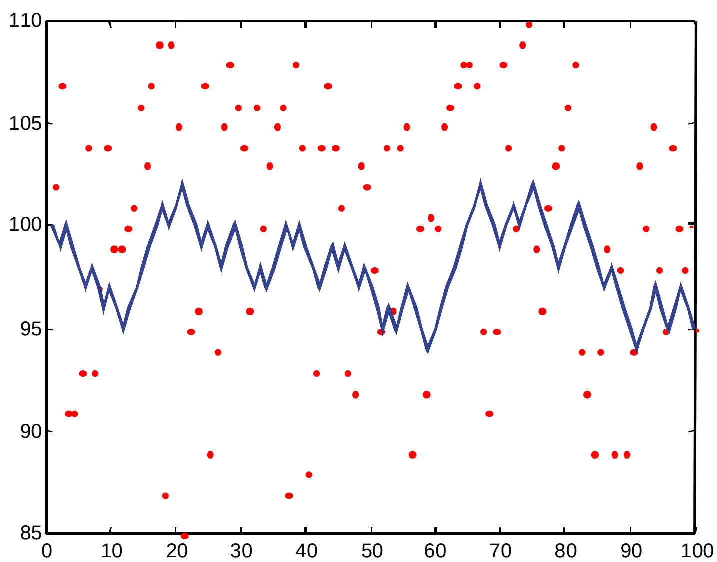
Time extraction:
1 Extrapolated leading edge
2 Pulse maximum

Problems:

ATWD reads out in reverse order



Pulse oscillations



Pulse background varies with stochastic component +-20

Reverse order algorithm

1 Check second halfe or last quarter of waveform for noise
calculating and comparing mean and max. Find pedestal
2 Look for pulses in the first part
8 extra history registers needed + some more - not too bad

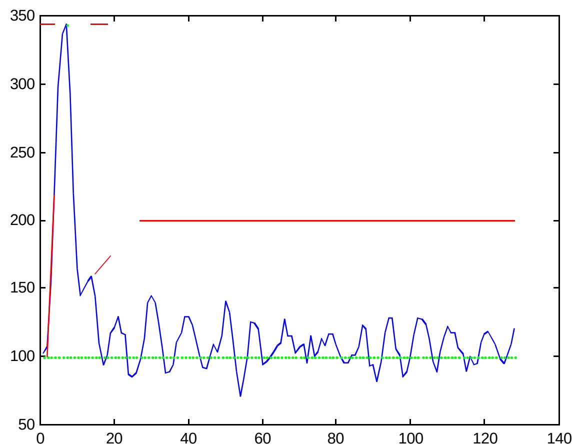
Better pedestal processing
Similar performance

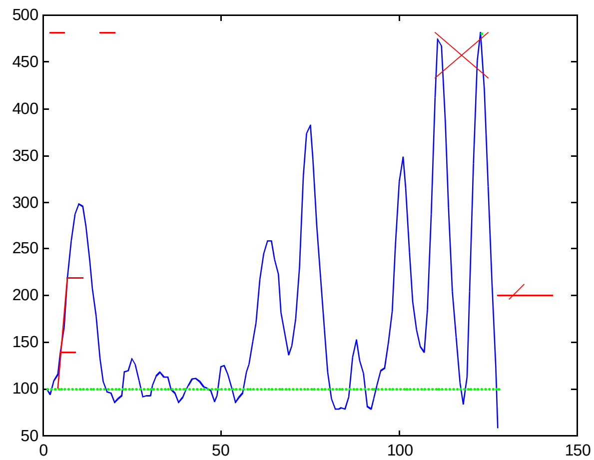
Evaluation

Detection evaluated by comparison with operator determined standard (tedious and not straight forward)

False negative 1 %

False positive 10 %





An approx timing resolution is obtained by comparing pulse width in a typical pulse (p) with a superimposed average pulse (a) where

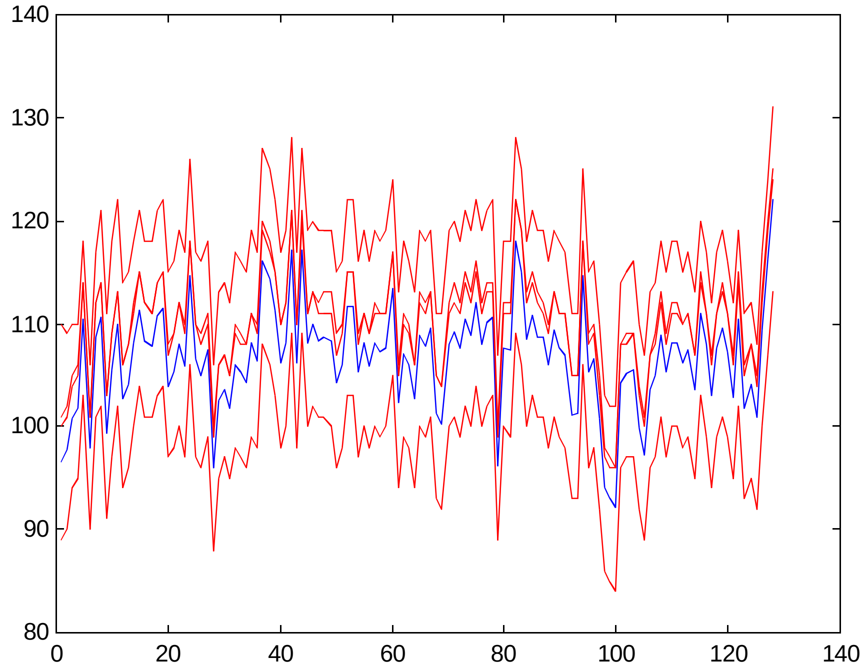
$$a = p * r$$

and r is the timing distribution

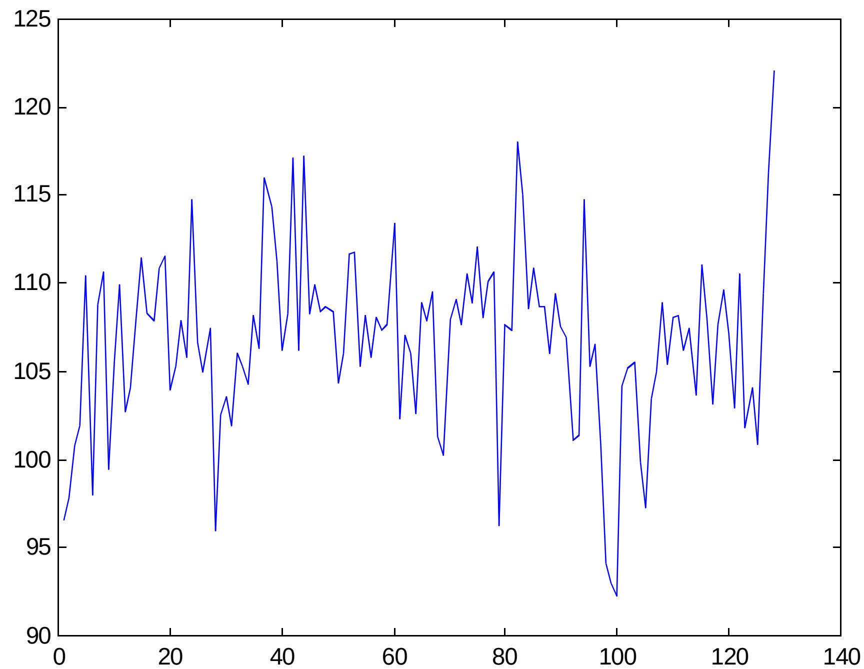
$$\mathrm{FWHM}(a)^2 \sim \mathrm{FWHM}(p)^2 + \mathrm{FWHM}(r)^2$$

FWHM(r) ~ 5 for ELET
0 for maximum

Evaluations from data difficult – simulations prefered where the true values are known
Pedestal data from Azriel – noise +- 10

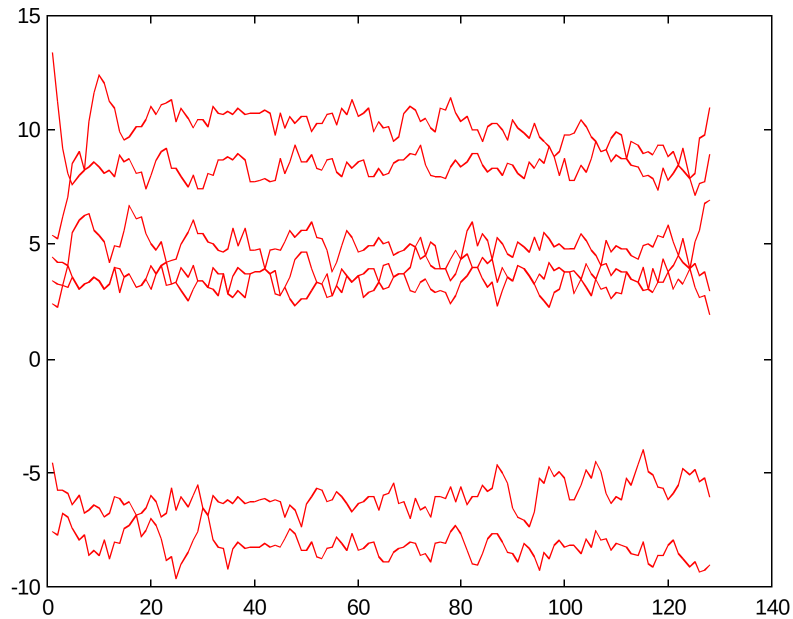


Apart from first frame alternating pedestal patterns similar

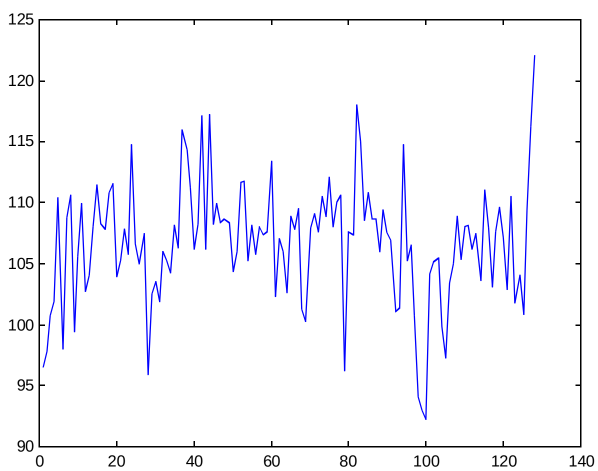


Average pattern

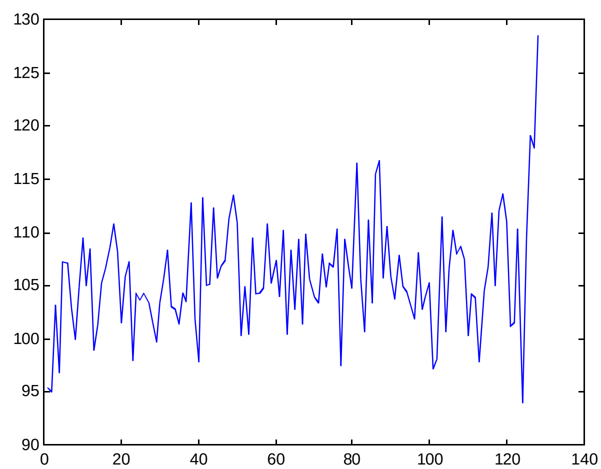
Remaining noise small +- 1



Stationary pattern characteristic from each ATWD

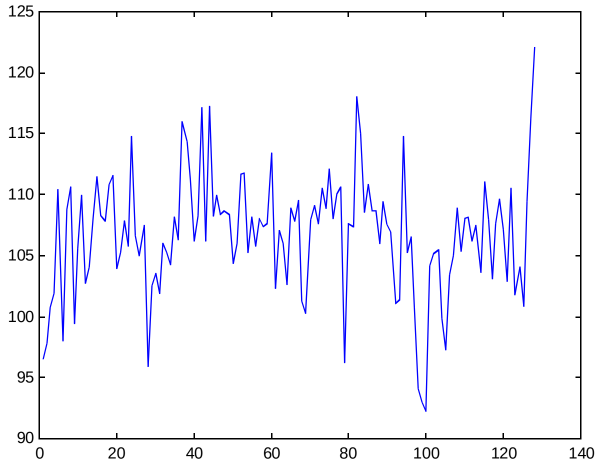


2,4,6,…

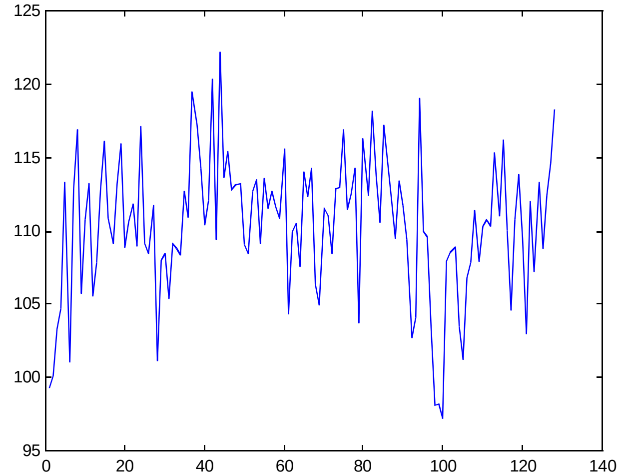


3,5,7,…

similar but not identical in different channels on the same ATWD

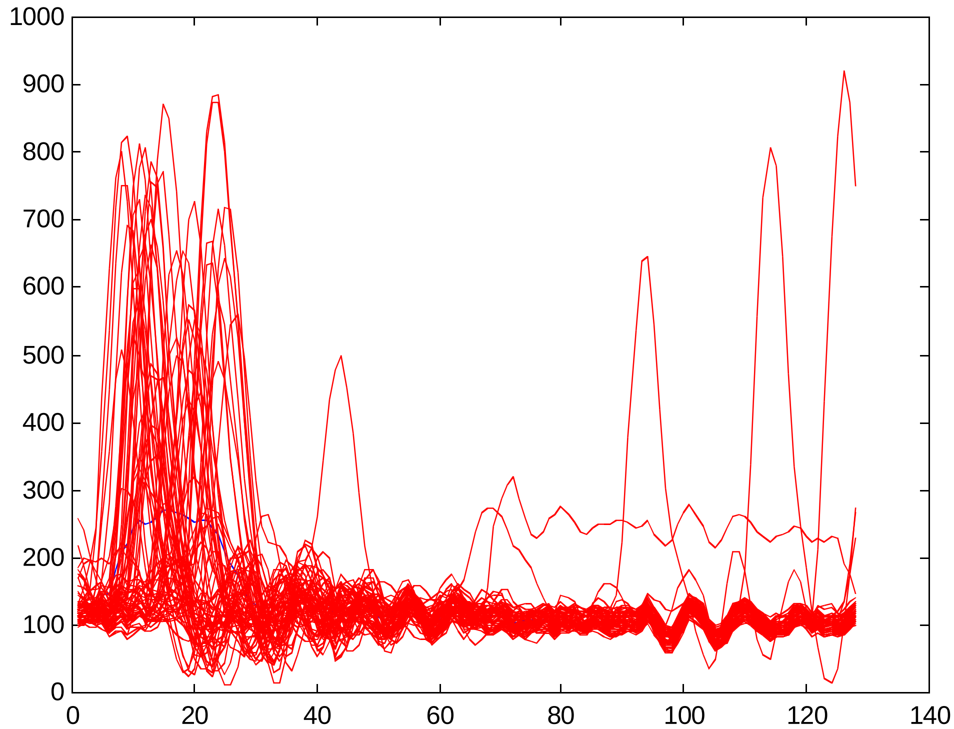


High gain

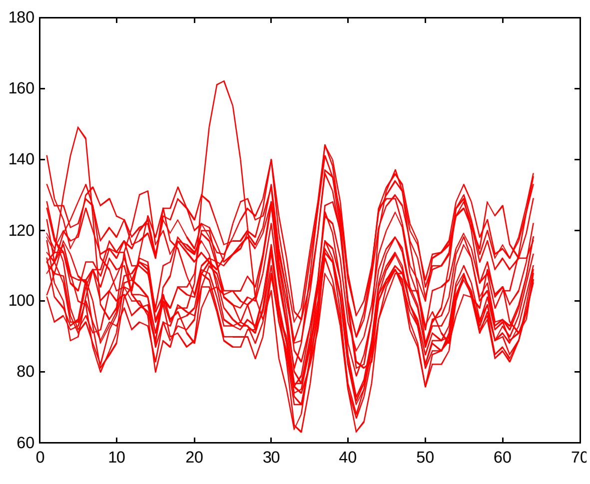


Low gain

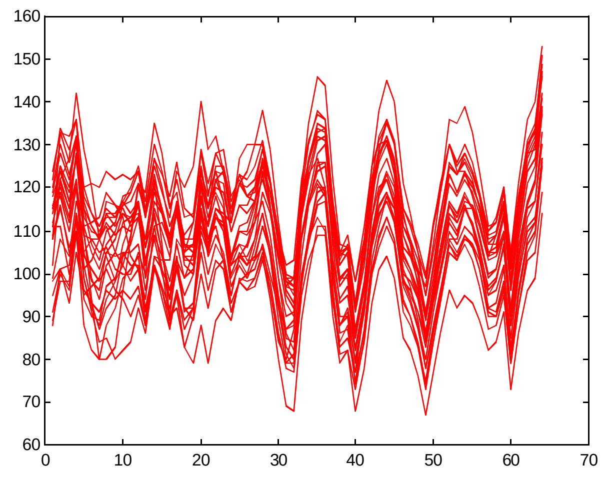
Pulses introduce larger stationary patterns



Full waveform for 2,4,6,…
Below the tails 65:248



2,4,6,…



3,5,7,…

Conclusions

There are many degrees of freedom that can be used to improve the feature extraction performance

Much can be done to improve both detection and timing performance

Trimming the algorithm parameters will improve the performance

We need to understand the noise sources better
Do we need to correct for some of them?

Reduced pulse oscillations will allow more ambitious shape control if we use a larger history memory

To-Do

Finish the S-18 FPGA design so that it can be tested in situ

Systematic study of noise and pulse shapes

Use simulated pulses to evaluate algorithms

trim algorithm parameters to optimize performance

Apply to pulses from the new DOM boards